WANLP 2022

The Seventh Arabic Natural Language Processing Workshop

Proceedings of the Workshop

December 8, 2022

The WANLP organizers gratefully acknowledge the support from the following sponsors.

Google Research

©2022 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-959429-27-2

Preface

أهلاً وسهلاً ومرحباً بكم! حيا كم في أبوظبي.

Ahlan wa-sahlan wa-marhaban bikum! Hayyaakum fi Abu Dhabi.
Hello and Welcome to Abu Dhabi!

Welcome to The Seventh Arabic Natural Language Processing Workshop (WANLP 2022) held with EMNLP 2022 in Abu Dhabi, UAE. Over the years, WANLP has developed a growing reputation as a high-quality venue for researchers and developers working on Arabic NLP, where they share and discuss their ongoing work. The first in the WANLP series was held in Doha, Qatar (EMNLP 2014), followed by Beijing, China (ACL 2015), Valencia, Spain (EACL 2017), Florence, Italy (ACL 2019), and in virtual mode in COLING 2020 and EACL 2021.

In this iteration of WANLP, we received 68 main workshop submissions (50 long, 14 short, and 4 demos). The total number of submissions is higher than all the earlier editions of the workshop. All papers submitted to the main workshop were reviewed by at least three reviewers each. Out of the 68 submissions, 36 were accepted: 31 long papers, two short papers, and three demo papers). We selected 13 papers for oral presentation and the rest as posters. We did not distinguish between long and short papers, or between oral and poster presentations in terms of quality.

WANLP 2022 included, for the first time, three shared tasks: The third edition of the Nuanced Arabic Dialect Identification (NADI) shared task, the Gender Rewriting Shared Task, and the Shared Task on Propaganda Detection in Arabic. NADI received submissions from 21 teams, 15 of which have system descriptions in the proceedings. The Gender Rewriting Shared Task received submissions from five teams, two of which have system descriptions in the proceedings. The Shared Task on Propaganda Detection in Arabic received submissions from 17 teams, 11 of which have system descriptions in the proceedings. The shared task system descriptions papers were reviewed by two reviewers each. Three additional shared task overview papers are included in the proceedings. The overview papers are presented in an oral session in the workshop.

For the second time, our workshop was able to secure sponsorship funding (Thanks to Google Research!) which we used to support student registrations.

In another success for the WANLP community, a new Special Interest Group on Arabic NLP (SIGARAB) was created in early 2022 by the advisory committee responsible for WANLP, building on its history of successful organization and collaboration.

Finally, we would like to thank everyone who submitted a paper to the workshop, as well as all the 83 members of the Program Committee, who worked hard to provide high-quality reviews on time. Organizing WANLP 2022 is a team effort.

Houda Bouamor, General Chair, on behalf of the workshop organizers.

Website of the workshop: http://wanlp2022.arabic-nlp.net/

Organizing Committee

General Chair

Houda Bouamor, Carnegie Mellon University in Qatar, Qatar

Program Chairs

Hend Al-Khalifa, King Saud University, KSA
Fethi Bougares, University of Le Mans, France
Kareem Darwish, aiXplain Inc.
Owen Rambow, Stony Brook University, USA

Publication Chairs

Ahmed Abdelali, Qatar Computing Research Institute, HBKU, Qatar
Nadi Tomeh, Université Sorbonne Paris Nord, France

Publicity Chairs

Salam Khalifa, Stony Brook University, USA
Wajdi Zaghouani, Hamad Bin Khalifa University, Qatar

Program Committee

Program Committee

Abdelmajid Ben-Hamadou, University of Sfax, Tunisia
AbdelRahim Elmadany, The University of British Columbia, Canada
Ahmed Abdelali, Qatar Computing Research Institute, HBKU, Qatar
Ahmed Ali, Qatar Computing Research Institute, HBKU, Qatar
Ahmed El Kholy, Columbia University, Microsoft, USA
Alexis Nasr, University of Marseille, France
Ali AlKhathlan, King Abdulaziz University, KSA
Almoataz B. Al-Said, Cairo University, Egypt
Aloulou Chafik, University of Sfax, Tunisia
Ann Bies, Linguistic Data Consortium, University of Pennsylvania, USA
Azzeddine Mazroui, University Mohamed I, Morocco
Bashar Alhafni, New York University in Abu Dhabi, UAE
Bashar Talafha, The University of British Columbia, Canada
Bassam Haddad, University of Petra, Jordan
Bayan AbuShawar, Al Ain University, UAE
Cal Peyser, Google Inc., USA
Chiyu Zhang, The University of British Columbia, Canada
Christian Khairallah, New York University in Abu Dhabi, UAE
Dana Abdulrahim, University of Bahrain, Bahrain
El Moatez Billah Nagoudi, The University of British Columbia, Canada
Fatemah Husain, Kuwait University, Kuwait
Fethi Bougares, Le Mans University, France
Firoj Alam, Qatar Computing Research Institute, HBKU, Qatar
Ganesh Jawahar, The University of British Columbia, Canada
Gilbert Badaro, American University of Beirut, Lebanon
Giovanni Da San Martino, University of Padova, Italy
Go Inoue, Mohamed Bin Zayed University of Artificial Intelligence, UAE
Hamada Nayel, Benha University, Egypt
Hamdy Mubarak, Qatar Computing Research Institute, HBKU, Qatar
Hend Al-Khalifa, King Saud University, KSA
Hossam Ahmed, Leiden Institute for Area Studies, Netherlands
Houda Bouamor, Carnegie Mellon University, Qatar
Ibrahim Abu Farha, University of Edinburgh, Scotland
Ife Adebara, The University of British Columbia, Canada
Imed Zitouni, Google Inc., USA
Imene Bensalem, Constantine 2 University, Algeria
Injy Hamed, New York University in Abu Dhabi, UAE
Jordan Kodner, Stony Brook University, USA
Kamel Smaili, LORIA, France
Kareem Darwish, aiXplain Inc., USA
Karim Bouzoubaa, Mohammad V University, Morocco
Khaled Shaalan, The British University in Dubai, UAE
Khaled Shaban, Qatar University, Qatar
Lamia Hadrich-Belguith, University of Sfax , Tunisia
Maram Hasanain, Qatar University, Qatar
Md Tawkat Islam Khondaker, The University of British Columbia, Canada

Mohamed Al-Badrashiny, aiXplain Inc., USA
Mohammad Abuoudeh, New York University in Abu Dhabi, UAE
Mourad Abbas, HCLA, Algeria
Muhammad Abdul-Mageed, The University of British Columbia, Canada
Mustafa Jarrar, Bir Zeit University, Palestine
Nada Ghneim, Higher Institute for Applied Sciences and Technology, Syria
Nada Almarwani, Taibah University, KSA
Nadi Tomeh, University Paris 13, France
Nizar Habash, New York University in Abu Dhabi, UAE
Nora Al-Twairesh, King Saud University, KSA
Omar Trigui, University of Sousse, Tunisia
Ossama Obeid, New York University in Abu Dhabi, UAE
Owen Rambow, Stony Brook University, USA
Peter Sullivan, The University of British Columbia, Canada
Preslav Nakov, Mohamed Bin Zayed University of Artificial Intelligence, UAE
Sahar Ghannay, LIMSI-CNRS, France
Sakhar Alkhereyf, King Abdulaziz City for Science and Technology, KSA
Salam Khalifa, Stony Brook University, USA
Salima Harrat, École Normale Supérieure (Bouzaréah), Algeria
Salima Mdhaffar, Le Mans University, France
Samah Aloufi, Taibah University, KSA
Samhaa R. El-Beltagy, Nile University, Egypt
Samia Touileb, University of Oslo, Norway
Seif Mechti, ISSEPS, Tunisia
Shady Elbassuoni, American University of Beirut, Lebanon
Shammur Absar Chowdhury, Qatar Computing Research Institute, HBKU, Qatar
Taha Zerrouki, Bouira University, Algeria
Tamer Elsayed, Qatar University, Qatar
Thamar Solorio, University of Houston, USA
Violetta Cavalli-Sforza, Al Akhawayn University, Morocco
Wajdi Zaghouani, Hamad Bin Khalifa University, Qatar
Wassim El-Hajj, American University of Beirut, Lebanon
Wei-Rui Chen, The University of British Columbia, Canada
Wissam Antoun, INRIA, France
Younes Samih, University of Düsseldorf, Germany
Yuval Marton, University of Washington, USA

Invited Speaker

Karim Bouzoubaa, Mohammadia School of Engineers at the Mohammed 5th University of Rabat

Keynote Talk: Digital Preservation of Arabic between Linguistics and AI

Karim Bouzoubaa
Mohammadia School of Engineers at the Mohammed 5th University of Rabat

Abstract: Languages are one of the oldest studied disciplines as they are intimately linked to the existence of human beings. The study of languages is a multidisciplinary field which has attracted the interest of several related fields such as linguistics and NLP, each providing additional knowledge for language understanding, learning, evolution, or preservation. From the technological point of view, computer science in general and artificial intelligence in particular study languages through natural language processing techniques, where the main goal is to discover linguistic patterns from corpora without resorting to linguists at all in many cases. Research in this field is diverse and currently benefits from advances in machine learning and deep learning techniques. One of the less studied aspects is the use and exploitation of these techniques for language preservation needs, for language comprehension needs or for the explanation of linguistic phenomena. The objective of this talk is to emphasize this perspective and to show through concrete cases how through the exploitation of several old and new computer and AI techniques, we can advance the digital preservation of Arabic and the explanation of some linguistic properties.

Bio: Karim Bouzoubaa is a Professor of computer science at the Mohammadia School of Engineers at the Mohammed 5th University of Rabat. Prof. Bouzoubaa holds a M.Sc. and a Ph.D. from Laval University in Canada in Artificial Intelligence and multi-agent systems fields. He is a research-driven professional with a distinctive combination of leadership, research and development, and education in the areas of Artificial Intelligence and Data Science. He contributed to the release of the Amine platform for the development of intelligent systems. He has published two books and over a hundred papers in top-ranked conferences and journals, taught at undergraduate and postgraduate levels, and worked on various R&D projects. He is the founding president of the Arabic Language Engineering Society in Morocco and the director of the Language Engineering lab. His research interests include Arabic NLP, NLP frameworks, Linguistic Resources and ontologies, IR and QA systems, Dialect processing, and Cognitive systems. He has led research teams, developed research programs, and has a long experience with most AI paradigms from the old to the newest ones. Prof Bouzoubaa is a Fulbright fellow and was a visiting professor at many institutions. He chaired and organized many international conferences, and was co-guest editor of the Special Issue on “Advances in Arabic Language Processing” for the International Journal on Information and Communication Technologies.

Table of Contents

CAraNER: The COVID-19 Arabic Named Entity Corpus
Abdulmohsen Al-Thubaity, Sakhar Alkhereyf, Wejdan Alzahrani and Alia Bahanshal 1

Joint Coreference Resolution for Zeros and non-Zeros in Arabic
Abdulrahman Aloraini, Sameer Pradhan and Massimo Poesio 11

SAIDS: A Novel Approach for Sentiment Analysis Informed of Dialect and Sarcasm
Abdelrahman Kaseb and Mona Farouk 22

AraBART: a Pretrained Arabic Sequence-to-Sequence Model for Abstractive Summarization
Moussa Kamal Eddine, Nadi Tomeh, Nizar Habash, Joseph Le Roux and Michalis Vazirgiannis 31

Towards Arabic Sentence Simplification via Classification and Generative Approaches
Nouran Khallaf, Serge Sharoff and Rasha Soliman 43

Generating Classical Arabic Poetry using Pre-trained Models
Nehal Elkaref, Mervat Abu-Elkheir, Maryam ElOraby and Mohamed Abdelgaber 53

A Benchmark Study of Contrastive Learning for Arabic Social Meaning
Md Tawkat Islam Khondaker, El Moatez Billah Nagoudi, AbdelRahim Elmadany, Muhammad Abdul-Mageed and Laks Lakshmanan, V.S. 63

Adversarial Text-to-Speech for low-resource languages
Ashraf Elneima and Mikołaj Bińkowski 76

NADI 2022: The Third Nuanced Arabic Dialect Identification Shared Task
Muhammad Abdul-Mageed, Chiyu Zhang, AbdelRahim Elmadany, Houda Bouamor and Nizar Habash 85

The Shared Task on Gender Rewriting
Bashar Alhafni, Nizar Habash, Houda Bouamor, Ossama Obeid, Sultan Alrowili, Daliyah AlZeer, Kawla Mohmad Shnqiti, Ahmed Elbakry, Muhammad ElNokrashy, Mohamed Gabr, Abderrahmane Issam, Abdelrahim Qaddoumi, Vijay Shanker and Mahmoud Zyate 98

Overview of the WANLP 2022 Shared Task on Propaganda Detection in Arabic
Firoj Alam, Hamdy Mubarak, Wajdi Zaghouani, Giovanni Da San Martino and Preslav Nakov 108

ArzEn-ST: A Three-way Speech Translation Corpus for Code-Switched Egyptian Arabic-English
Injy Hamed, Nizar Habash, Slim Abdennadher and Ngoc Thang Vu 119

Maknuune: A Large Open Palestinian Arabic Lexicon
Shahd Salah Uddin Dibas, Christian Khairallah, Nizar Habash, Omar Fayez Sadi, Tariq Sairafy, Karmel Sarabta and Abrar Ardah 131

Developing a Tag-Set and Extracting the Morphological Lexicons to Build a Morphological Analyzer for Egyptian Arabic
Amany Fashwan and Sameh Alansary 142

A Weak Supervised Transfer Learning Approach for Sentiment Analysis to the Kuwaiti Dialect
Fatemah Husain, Hana Al-Ostad and Halima Omar 161

Mawqif: A Multi-label Arabic Dataset for Target-specific Stance Detection
Nora Saleh Alturayeif, Hamzah Abdullah Luqman and Moataz Aly Kamaleldin Ahmed 174

Assessing the Linguistic Knowledge in Arabic Pre-trained Language Models Using Minimal Pairs
Wafa Abdullah Alrajhi, Hend Al-Khalifa and Abdulmalik AlSalman 185

Identifying Code-switching in Arabizi
Safaa Shehadi and Shuly Wintner 194

Authorship Verification for Arabic Short Texts Using Arabic Knowledge-Base Model (AraKB)
Fatimah Alqahtani and Helen Yannakoudakis 205

A Semi-supervised Approach for a Better Translation of Sentiment in Dialectical Arabic UGT
Hadeel Saadany, Constantin Orăsan, Emad Mohamed and Ashraf Tantawy 214

Cross-lingual transfer for low-resource Arabic language understanding
Khadige Abboud, Olga Golovneva and Christopher DiPersio 225

Improving POS Tagging for Arabic Dialects on Out-of-Domain Texts
Noor Abo Mokh, Daniel Dakota and Sandra Kübler 238

Domain Adaptation for Arabic Crisis Response
Reem Alrashdi and Simon O'Keefe 249

Weakly and Semi-Supervised Learning for Arabic Text Classification using Monodialectal Language Models
Reem AlYami and Rabah Al-Zaidy 260

Event-Based Knowledge MLM for Arabic Event Detection
Asma Z Yamani, Amjad K Alsulami and Rabeah A Al-Zaidy 273

Establishing a Baseline for Arabic Patents Classification: A Comparison of Twelve Approaches
Taif Omar Al-Omar, Hend Al-Khalifa and Rawan Al-Matham 287

Towards Learning Arabic Morphophonology
Salam Khalifa, Jordan Kodner and Owen Rambow 295

AraDepSu: Detecting Depression and Suicidal Ideation in Arabic Tweets Using Transformers
Mariam Hassib, Nancy Hossam, Jolie Sameh and Marwan Torki 302

End-to-End Speech Translation of Arabic to English Broadcast News
Fethi Bougares and Salim Jouili 312

Arabic Keyphrase Extraction: Enhancing Deep Learning Models with Pre-trained Contextual Embedding and External Features
Randah Alharbi and Husni Al-Muhtasab 320

ArabIE: Joint Entity, Relation and Event Extraction for Arabic
Niama El Khbir, Nadi Tomeh and Thierry Charnois 331

Emoji Sentiment Roles for Sentiment Analysis: A Case Study in Arabic Texts
Shatha Ali A. Hakami, Robert Hendley and Phillip Smith 346

Gulf Arabic Diacritization: Guidelines, Initial Dataset, and Results
nouf alabbasi, Mohamed Al-Badrashiny, Maryam Aldahmani, Ahmed AlDhanhani, Abdullah Saleh Alhashmi, Fawaghy Ahmed Alhashmi, Khalid Al Hashemi, Rama Emad Alkhobbi, Shamma T Al Maazmi, Mohammed Ali Alyafeai, Mariam M Alzaabi, Mohamed Saqer Alzaabi, Fatma Khalid Badri, Kareem Darwish, Ehab Mansour Diab, Muhammad Morsy Elmallah, Amira Ayman Elnashar, Ashraf Hatim Elneima, MHD Tameem Kabbani, Nour Rabih, Ahmad Saad and Ammar Mamoun Sousou . 356

Learning From Arabic Corpora But Not Always From Arabic Speakers: A Case Study of the Arabic Wikipedia Editions
Saied Alshahrani, Esma Wali and Jeanna Matthews 361

A Pilot Study on the Collection and Computational Analysis of Linguistic Differences Amongst Men and Women in a Kuwaiti Arabic WhatsApp Dataset
Hesah Aldihan, Robert Gaizauskas and Susan Fitzmaurice 372

Beyond Arabic: Software for Perso-Arabic Script Manipulation
Alexander Gutkin, Cibu Johny, Raiomond Doctor, Brian Roark and Richard Sproat 381

Coreference Annotation of an Arabic Corpus using a Virtual World Game
Wateen Abdullah Aliady, Abdulrahman Aloraini, Christopher Madge, Juntao Yu, Richard Bartle and Massimo Poesio 388

NatiQ: An End-to-end Text-to-Speech System for Arabic
Ahmed Abdelali, Nadir Durrani, Cenk Demiroglu, Fahim Dalvi, Hamdy Mubarak and Kareem Darwish 394

The Effect of Arabic Dialect Familiarity on Data Annotation
Ibrahim Abu Farha and Walid Magdy 399

Optimizing Naive Bayes for Arabic Dialect Identification
Tommi Jauhiainen, Heidi Jauhiainen and Krister Lindén 409

iCompass Working Notes for the Nuanced Arabic Dialect Identification Shared task
Abir Messaoudi, Chayma Fourati, Hatem Haddad and Moez BenHajhmida 415

TF-IDF or Transformers for Arabic Dialect Identification? ITFLOWS participation in the NADI 2022 Shared Task
Fouad Shammary, Yiyi Chen, Zsolt T Kardkovacs, Mehwish Alam and Haithem Afli 420

Domain-Adapted BERT-based Models for Nuanced Arabic Dialect Identification and Tweet Sentiment Analysis
Giyaseddin Bayrak and ABDUL MAJEED ISSIFU 425

Benchmarking transfer learning approaches for sentiment analysis of Arabic dialect
emna fsih, Sameh Kchaou, Rahma Boujelbane and Lamia Hadrich-Belguith 431

SQU-CS @ NADI 2022: Dialectal Arabic Identification using One-vs-One Classification with TF-IDF Weights Computed on Character n-grams
Abdulrahman Khalifa AAlAbdulsalam 436

Ahmed and Khalil at NADI 2022: Transfer Learning and Addressing Class Imbalance for Arabic Dialect Identification and Sentiment Analysis
Ahmed Oumar and Khalil Mrini 442

Arabic Sentiment Analysis by Pretrained Ensemble
Abdelrahim Qaddoumi 447

Dialect & Sentiment Identification in Nuanced Arabic Tweets Using an Ensemble of Prompt-based, Fine-tuned, and Multitask BERT-Based Models
Reem Abdel-Salam 452

On The Arabic Dialects' Identification: Overcoming Challenges of Geographical Similarities Between Arabic dialects and Imbalanced Datasets
Salma Jamal, Aly M .Kassem, Omar Mohamed and Ali Ashraf 458

Arabic dialect identification using machine learning and transformer-based models: Submission to the NADI 2022 Shared Task
Nouf AlShenaifi and Aqil Azmi 464

NLP DI at NADI Shared Task Subtask-1: Sub-word Level Convolutional Neural Models and Pre-trained Binary Classifiers for Dialect Identification
Vani Kanjirangat, Tanja Samardzic, Ljiljana Dolamic and Fabio Rinaldi 468

Word Representation Models for Arabic Dialect Identification
Mahmoud Sobhy, Ahmed H. Abu El-Atta, Ahmed A. El-Sawy and Hamada Nayel 474

Building an Ensemble of Transformer Models for Arabic Dialect Classification and Sentiment Analysis
Abdullah Salem Khered, Ingy Yasser Hassan Abdou Abdelhalim and Riza Batista-Navarro . . 479

Arabic Dialect Identification and Sentiment Classification using Transformer-based Models
Joseph Attieh and Fadi Hassan 485

Generative Approach for Gender-Rewriting Task with ArabicT5
Sultan Alrowili and Vijay Shanker 491

AraProp at WANLP 2022 Shared Task: Leveraging Pre-Trained Language Models for Arabic Propaganda Detection
Gaurav Singh 496

TUB at WANLP22 Shared Task: Using Semantic Similarity for Propaganda Detection in Arabic
Salar Mohtaj and Sebastian Möller 501

SI2M & AIOX Labs at WANLP 2022 Shared Task: Propaganda Detection in Arabic, A Data Augmentation and Name Entity Recognition Approach
Kamel Gaanoun and Imade Benelallam 506

iCompass at WANLP 2022 Shared Task: ARBERT and MARBERT for Multilabel Propaganda Classification of Arabic Tweets
Bilel - Taboubi, Bechir Brahem and Hatem Haddad 511

ChavanKane at WANLP 2022 Shared Task: Large Language Models for Multi-label Propaganda Detection
Tanmay Chavan and Aditya Manish Kane 515

AraBERT Model for Propaganda Detection
Mohamad Sharara, Wissam Mohamad, Ralph Tawil, Ralph Chobok, Wolf Assi and Antonio Tannoury 520

AraBEM at WANLP 2022 Shared Task: Propaganda Detection in Arabic Tweets
Eshrag Ali Refaee, Basem Ahmed and Motaz Saad 524

IITD at WANLP 2022 Shared Task: Multilingual Multi-Granularity Network for Propaganda Detection
Shubham Mittal and Preslav Nakov 529

Pythoneers at WANLP 2022 Shared Task: Monolingual AraBERT for Arabic Propaganda Detection and Span Extraction
Joseph Attieh and Fadi Hassan 534

CNLP-NITS-PP at WANLP 2022 Shared Task: Propaganda Detection in Arabic using Data Augmentation and AraBERT Pre-trained Model
Sahinur Rahman Laskar, Rahul Singh, Abdullah Faiz Ur Rahman Khilji, Riyanka Manna, Partha Pakray and Sivaji Bandyopadhyay 541

NGU CNLP atWANLP 2022 Shared Task: Propaganda Detection in Arabic
Ahmed Samir Hussein, Abu Bakr Soliman Mohammad, Mohamed Ibrahim, Laila Hesham Afify and Samhaa R. El-Beltagy 545

Program

Thursday, December 8, 2022

09:00 - 09:05 *Opening Remarks*

09:10 - 10:00 *Invited Talk*

10:00 - 10:30 *Session 1 - Information Extraction (in-Person)*

CAraNER: The COVID-19 Arabic Named Entity Corpus
Abdulmohsen Al-Thubaity, Sakhar Alkhereyf, Wejdan Alzahrani and Alia Bahanshal

Joint Coreference Resolution for Zeros and non-Zeros in Arabic
Abdulrahman Aloraini, Sameer Pradhan and Massimo Poesio

10:30 - 11:00 *Coffee Break*

11:00 - 12:30 *Session 2 - NLU/NLG (in-Person)*

SAIDS: A Novel Approach for Sentiment Analysis Informed of Dialect and Sarcasm
Abdelrahman Kaseb and Mona Farouk

AraBART: a Pretrained Arabic Sequence-to-Sequence Model for Abstractive Summarization
Moussa Kamal Eddine, Nadi Tomeh, Nizar Habash, Joseph Le Roux and Michalis Vazirgiannis

Towards Arabic Sentence Simplification via Classification and Generative Approaches
Nouran Khallaf, Serge Sharoff and Rasha Soliman

Generating Classical Arabic Poetry using Pre-trained Models
Nehal Elkaref, Mervat Abu-Elkheir, Maryam ElOraby and Mohamed Abdelgaber

A Benchmark Study of Contrastive Learning for Arabic Social Meaning
Md Tawkat Islam Khondaker, El Moatez Billah Nagoudi, AbdelRahim Elmadany, Muhammad Abdul-Mageed and Laks Lakshmanan, V.S.

Adversarial Text-to-Speech for low-resource languages
Ashraf Elneima and Mikołaj Bińkowski

Thursday, December 8, 2022 (continued)

12:30 - 14:00 *Lunch Break*

14:00 - 14:45 *Panel discussion: Young Researchers in Arabic NLP*

14:45 - 15:00 *Best Paper Award Oral Presentation (in-Person)*

15:00 - 15:30 *Shared Task Papers (in-Person)*

NADI 2022: The Third Nuanced Arabic Dialect Identification Shared Task
Muhammad Abdul-Mageed, Chiyu Zhang, AbdelRahim Elmadany, Houda Bouamor and Nizar Habash

The Shared Task on Gender Rewriting
Bashar Alhafni, Nizar Habash, Houda Bouamor, Ossama Obeid, Sultan Alrowili, Daliyah AlZeer, Kawla Mohmad Shnqiti, Ahmed Elbakry, Muhammad El-Nokrashy, Mohamed Gabr, Abderrahmane Issam, Abdelrahim Qaddoumi, Vijay Shanker and Mahmoud Zyate

Overview of the WANLP 2022 Shared Task on Propaganda Detection in Arabic
Firoj Alam, Hamdy Mubarak, Wajdi Zaghouani, Giovanni Da San Martino and Preslav Nakov

15:30 - 16:00 *Coffee Break*

16:00 - 17:00 *Session 3 - Arabic Dialects (in-Person)*

ArzEn-ST: A Three-way Speech Translation Corpus for Code-Switched Egyptian Arabic-English
Injy Hamed, Nizar Habash, Slim Abdennadher and Ngoc Thang Vu

Maknuune: A Large Open Palestinian Arabic Lexicon
Shahd Salah Uddin Dibas, Christian Khairallah, Nizar Habash, Omar Fayez Sadi, Tariq Sairafy, Karmel Sarabta and Abrar Ardah

Developing a Tag-Set and Extracting the Morphological Lexicons to Build a Morphological Analyzer for Egyptian Arabic
Amany Fashwan and Sameh Alansary

A Weak Supervised Transfer Learning Approach for Sentiment Analysis to the Kuwaiti Dialect
Fatemah Husain, Hana Al-Ostad and Halima Omar

Thursday, December 8, 2022 (continued)

17:00 - 18:15 *Main Workshop Posters (in-Person & virtual)*

Mawqif: A Multi-label Arabic Dataset for Target-specific Stance Detection
Nora Saleh Alturayeif, Hamzah Abdullah Luqman and Moataz Aly Kamaleldin Ahmed

Assessing the Linguistic Knowledge in Arabic Pre-trained Language Models Using Minimal Pairs
Wafa Abdullah Alrajhi, Hend Al-Khalifa and Abdulmalik AlSalman

Identifying Code-switching in Arabizi
Safaa Shehadi and Shuly Wintner

Authorship Verification for Arabic Short Texts Using Arabic Knowledge-Base Model (AraKB)
Fatimah Alqahtani and Helen Yannakoudakis

A Semi-supervised Approach for a Better Translation of Sentiment in Dialectical Arabic UGT
Hadeel Saadany, Constantin Orăsan, Emad Mohamed and Ashraf Tantawy

Cross-lingual transfer for low-resource Arabic language understanding
Khadige Abboud, Olga Golovneva and Christopher DiPersio

Improving POS Tagging for Arabic Dialects on Out-of-Domain Texts
Noor Abo Mokh, Daniel Dakota and Sandra Kübler

Domain Adaptation for Arabic Crisis Response
Reem Alrashdi and Simon O'Keefe

Weakly and Semi-Supervised Learning for Arabic Text Classification using Monodialectal Language Models
Reem AlYami and Rabah Al-Zaidy

Event-Based Knowledge MLM for Arabic Event Detection
Asma Z Yamani, Amjad K Alsulami and Rabeah A Al-Zaidy

Establishing a Baseline for Arabic Patents Classification: A Comparison of Twelve Approaches
Taif Omar Al-Omar, Hend Al-Khalifa and Rawan Al-Matham

Towards Learning Arabic Morphophonology
Salam Khalifa, Jordan Kodner and Owen Rambow

AraDepSu: Detecting Depression and Suicidal Ideation in Arabic Tweets Using Transformers
Mariam Hassib, Nancy Hossam, Jolie Sameh and Marwan Torki

End-to-End Speech Translation of Arabic to English Broadcast News
Fethi Bougares and Salim Jouili

Arabic Keyphrase Extraction: Enhancing Deep Learning Models with Pre-trained Contextual Embedding and External Features
Randah Alharbi and Husni Al-Muhtasab

ArabIE: Joint Entity, Relation and Event Extraction for Arabic
Niama El Khbir, Nadi Tomeh and Thierry Charnois

Emoji Sentiment Roles for Sentiment Analysis: A Case Study in Arabic Texts
Shatha Ali A. Hakami, Robert Hendley and Phillip Smith

Gulf Arabic Diacritization: Guidelines, Initial Dataset, and Results
nouf alabbasi, Mohamed Al-Badrashiny, Maryam Aldahmani, Ahmed AlDhanhani, Abdullah Saleh Alhashmi, Fawaghy Ahmed Alhashmi, Khalid Al Hashemi, Rama Emad Alkhobbi, Shamma T Al Maazmi, Mohammed Ali Alyafeai, Mariam M Alzaabi, Mohamed Saqer Alzaabi, Fatma Khalid Badri, Kareem Darwish, Ehab Mansour Diab, Muhammad Morsy Elmallah, Amira Ayman Elnashar, Ashraf Hatim Elneima, MHD Tameem Kabbani, Nour Rabih, Ahmad Saad and Ammar Mamoun Sousou

Learning From Arabic Corpora But Not Always From Arabic Speakers: A Case Study of the Arabic Wikipedia Editions
Saied Alshahrani, Esma Wali and Jeanna Matthews

A Pilot Study on the Collection and Computational Analysis of Linguistic Differences Amongst Men and Women in a Kuwaiti Arabic WhatsApp Dataset
Hesah Aldihan, Robert Gaizauskas and Susan Fitzmaurice

17:00 - 18:15 *EMNLP Findings Posters (in-Person & virtual)*

Improving English-Arabic Transliteration with Phonemic Memories
Yan Song, Shengyi JIANG, Lianxi Wang, Xiangyu Pang, Renze Lou and Yuanhe Tian

Sarcasm Detection is Way Too Easy! An Empirical Comparison of Human and Machine Sarcasm Detection
Walid Magdy, Silviu Oprea, Steven Wilson and Ibrahim Abu Farha

17:00 - 18:15 *Demos (in-Person & virtual)*

Beyond Arabic: Software for Perso-Arabic Script Manipulation
Alexander Gutkin, Cibu Johny, Raiomond Doctor, Brian Roark and Richard Sproat

Coreference Annotation of an Arabic Corpus using a Virtual World Game
Wateen Abdullah Aliady, Abdulrahman Aloraini, Christopher Madge, Juntao Yu, Richard Bartle and Massimo Poesio

NatiQ: An End-to-end Text-to-Speech System for Arabic
Ahmed Abdelali, Nadir Durrani, Cenk Demiroglu, Fahim Dalvi, Hamdy Mubarak and Kareem Darwish

17:00 - 18:15 *NADI Shared Task (in-Person & virtual)*

Optimizing Naive Bayes for Arabic Dialect Identification
Tommi Jauhiainen, Heidi Jauhiainen and Krister Lindén

iCompass Working Notes for the Nuanced Arabic Dialect Identification Shared task
Abir Messaoudi, Chayma Fourati, Hatem Haddad and Moez BenHajhmida

TF-IDF or Transformers for Arabic Dialect Identification? ITFLOWS participation in the NADI 2022 Shared Task
Fouad Shammary, Yiyi Chen, Zsolt T Kardkovacs, Mehwish Alam and Haithem Afli

Domain-Adapted BERT-based Models for Nuanced Arabic Dialect Identification and Tweet Sentiment Analysis
Giyaseddin Bayrak and ABDUL MAJEED ISSIFU

Benchmarking transfer learning approaches for sentiment analysis of Arabic dialect
emna fsih, Sameh Kchaou, Rahma Boujelbane and Lamia Hadrich-Belguith

SQU-CS @ NADI 2022: Dialectal Arabic Identification using One-vs-One Classification with TF-IDF Weights Computed on Character n-grams
Abdulrahman Khalifa AAlAbdulsalam

Ahmed and Khalil at NADI 2022: Transfer Learning and Addressing Class Imbalance for Arabic Dialect Identification and Sentiment Analysis
Ahmed Oumar and Khalil Mrini

Arabic Sentiment Analysis by Pretrained Ensemble
Abdelrahim Qaddoumi

Dialect & Sentiment Identification in Nuanced Arabic Tweets Using an Ensemble of Prompt-based, Fine-tuned, and Multitask BERT-Based Models
Reem Abdel-Salam

On The Arabic Dialects' Identification: Overcoming Challenges of Geographical Similarities Between Arabic dialects and Imbalanced Datasets
Salma Jamal, Aly M .Kassem, Omar Mohamed and Ali Ashraf

Arabic dialect identification using machine learning and transformer-based models: Submission to the NADI 2022 Shared Task
Nouf AlShenaifi and Aqil Azmi

NLP DI at NADI Shared Task Subtask-1: Sub-word Level Convolutional Neural Models and Pre-trained Binary Classifiers for Dialect Identification
Vani Kanjirangat, Tanja Samardzic, Ljiljana Dolamic and Fabio Rinaldi

Word Representation Models for Arabic Dialect Identification
Mahmoud Sobhy, Ahmed H. Abu El-Atta, Ahmed A. El-Sawy and Hamada Nayel

Building an Ensemble of Transformer Models for Arabic Dialect Classification and Sentiment Analysis
Abdullah Salem Khered, Ingy Yasser Hassan Abdou Abdelhalim and Riza Batista-Navarro

Arabic Dialect Identification and Sentiment Classification using Transformer-based Models
Joseph Attieh and Fadi Hassan

17:00 - 18:15 *Gender Rewriting Shared Task Posters (in-Person & virtual)*

Thursday, December 8, 2022 (continued)

Generative Approach for Gender-Rewriting Task with ArabicT5
Sultan Alrowili and Vijay Shanker

17:00 - 18:15 *Propaganda Detection Shared Task Posters (in-Person & virtual)*

AraProp at WANLP 2022 Shared Task: Leveraging Pre-Trained Language Models for Arabic Propaganda Detection
Gaurav Singh

TUB at WANLP22 Shared Task: Using Semantic Similarity for Propaganda Detection in Arabic
Salar Mohtaj and Sebastian Möller

SI2M & AIOX Labs at WANLP 2022 Shared Task: Propaganda Detection in Arabic, A Data Augmentation and Name Entity Recognition Approach
Kamel Gaanoun and Imade Benelallam

iCompass at WANLP 2022 Shared Task: ARBERT and MARBERT for Multilabel Propaganda Classification of Arabic Tweets
Bilel - Taboubi, Bechir Brahem and Hatem Haddad

ChavanKane at WANLP 2022 Shared Task: Large Language Models for Multi-label Propaganda Detection
Tanmay Chavan and Aditya Manish Kane

AraBERT Model for Propaganda Detection
Mohamad Sharara, Wissam Mohamad, Ralph Tawil, Ralph Chobok, Wolf Assi and Antonio Tannoury

AraBEM at WANLP 2022 Shared Task: Propaganda Detection in Arabic Tweets
Eshrag Ali Refaee, Basem Ahmed and Motaz Saad

IITD at WANLP 2022 Shared Task: Multilingual Multi-Granularity Network for Propaganda Detection
Shubham Mittal and Preslav Nakov

Pythoneers at WANLP 2022 Shared Task: Monolingual AraBERT for Arabic Propaganda Detection and Span Extraction
Joseph Attieh and Fadi Hassan

Thursday, December 8, 2022 (continued)

CNLP-NITS-PP at WANLP 2022 Shared Task: Propaganda Detection in Arabic using Data Augmentation and AraBERT Pre-trained Model
Sahinur Rahman Laskar, Rahul Singh, Abdullah Faiz Ur Rahman Khilji, Riyanka Manna, Partha Pakray and Sivaji Bandyopadhyay

NGU CNLP atWANLP 2022 Shared Task: Propaganda Detection in Arabic
Ahmed Samir Hussein, Abu Bakr Soliman Mohammad, Mohamed Ibrahim, Laila Hesham Afify and Samhaa R. El-Beltagy

17:00 - 18:30 *Closing Ceremony*